
 M E S S A G E
 (TW)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_lp protein - protein database search, using Smith-Waterman algorithm
 Run on: Sat May 13 08:08:38 2000; Maspar time 23.60 Seconds
 Tabular output not generated. 668.443 Million cell updates/sec

Title: >US-09-331-631-3
 Description: (1-666) from US09331631.pep (1 of 5)
 Perfect Score: 4913
 Sequence: 1 MAINTSNLCSLFLSLFL.....SSRSRKQQLVSLIDVGF 666

Scoring table:
 Gap 11
 PAM 150

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: a-geneseg35
 1.genesegp

Statistics: Mean 37.214; Variance 174.766; scale 0.213

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	4913	100.0	666	1	Macadamia integrifolia	0.00e+00
2	4746	96.6	666	1	Macadamia integrifolia	0.00e+00
3	4496	91.5	625	1	Macadamia integrifolia	0.00e+00
4	1353	27.5	566	1	Sequence encoded by 67	3.64e-107
5	1236	25.2	590	1	Gossypium hirsutum ant	1.29e-96
6	1193	22.3	523	1	Theobroma cacao antimi	9.54e-93
7	1106	22.5	593	1	Zea mays antimicrobial	6.29e-85
8	1036	21.1	489	1	G. max SBP2 protein.	1.19e-78
9	998	20.3	524	1	G. max SBP2 protein.	3.01e-75
10	981	20.0	605	1	Glycine max antimicrob	9.98e-74
11	933	19.0	444	1	G. max truncated SBP2	1.94e-69
12	933	19.0	626	1	peanut allergen Ara hi	1.94e-69
13	922	18.8	614	1	peanut allergen Ara hi	1.86e-68
14	922	18.8	614	1	Arachis hypogaea antim	1.86e-68
15	890	18.1	409	1	G. max truncated SBP2	1.33e-65
16	842	17.1	637	1	Hordeum vulgare anti	2.49e-61
17	180	3.7	395	1	Mouse SRY-related prot	1.10e-04
18	174	3.5	303	1	Diofililaria limulids pa	3.15e-04
19	145	3.0	186	1	Trypanosoma cruzi anti	4.48e-02
20	147	3.0	567	1	Polypeptide fragment e	3.20e-02
21	136	2.8	432	1	Human regulatory molec	1.99e-01
22	130	2.6	265	1	HTIV-1 protein express	5.28e-01
23	127	2.6	316	1	P.falciparum USA-R-NR	8.58e-01

ID	Score	Query Match	Length	ID	Description	Pred. No.
24	127	2.6	493	1	R26944	P.falciparum USA gene
25	130	2.6	541	1	W37148	Mammalian Ena (Mena).
26	126	2.6	562	1	R70491	Leucocytozoan protozoa
27	130	2.6	783	1	W37151	Mouse neural Mena+ pro
28	130	2.6	787	1	W37152	Mouse neural Mena+ pr
29	130	2.6	802	1	W37153	Mouse neural Mena+ p
30	129	2.6	1529	1	R37985	CORK potassium channel
31	125	2.5	1382	1	W31867	Human metastasis-assoc
32	116	2.4	190	1	R82664	E. coli RNase E.
33	116	2.4	351	1	W44681	Escherichia coli RNase
34	120	2.4	351	1	R24393	Sequence of Histidine
35	116	2.4	436	1	W03652	Human 70k U1 snRNP pro
36	118	2.4	521	1	W74802	Human secreted protein
37	116	2.4	614	1	R82630	70k autoantigen, part
38	117	2.4	1178	1	W30763	Mannose-1-phosphate tr
39	111	2.3	28	1	W62841	Stenocarpus sinuatus a
40	113	2.3	167	1	W89768	Staphylococcus aureus
41	112	2.3	325	1	W59645	Amino acid sequence of
42	114	2.3	1311	1	W72971	Precis coenia patched
43	114	2.3	1311	1	W32197	Precis coenia (butlerf
44	112	2.3	1343	1	W31866	Mouse metastasis-assoc
45	110	2.2	482	1	Y07067	Renal cancer associate

ALIGNMENTS

RESULT	ID	Score	Query Match	Length	ID	Description	Pred. No.
1	W62829	100.0%	Score 4913; DB 1; Length 666;				
2	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
3	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
4	W62829	100.0%	Score 4913; DB 1; Length 666;				
5	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
6	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
7	W62829	100.0%	Score 4913; DB 1; Length 666;				
8	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
9	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
10	W62829	100.0%	Score 4913; DB 1; Length 666;				
11	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
12	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
13	W62829	100.0%	Score 4913; DB 1; Length 666;				
14	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
15	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
16	W62829	100.0%	Score 4913; DB 1; Length 666;				
17	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
18	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
19	W62829	100.0%	Score 4913; DB 1; Length 666;				
20	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
21	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
22	W62829	100.0%	Score 4913; DB 1; Length 666;				
23	W62829	100.0%	Best Local Similarity 100.0%; Pred. No. 0.00e+00;				
24	W62829	100.0%	Matches 666; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				

D	b	241	EKGKOSDNYYVDERLSLSTFPTKEEGHISVLNPFGRSKLLPALKNYRLVYLEAPNA	3000
O	y	241	EEGEEKOSDNYIYDERLSLSTFPTKEEGHISVLNPFGRSKLLPALKNYRLVYLEAPNA	3000
D	b	301	FVLPLTHLDADAILLVLTGGRGALAKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNERTL	3600
O	y	301	FVLPLTHLDADAILLVLTGGRGALAKMIHRDNRESYNLECGDVIRIPAGTTFYLINRDNERTL	3600
D	b	361	HIAFLQITSPRGQYKEFFPAAGGNPPERYLSTEFEKELLEALNTQAARLGVLGGQREGV	4200
O	y	361	HIAFLQITSPRGQYKEFFPAAGGNPPERYLSTEFEKELLEALNTQAARLGVLGGQREGV	4200
D	b	421	IISASQEQIRLRTDDSSRRMWHIRRGESSRCGPYNLFNKRPILSNKYGOAYEVPEDYR	4800
O	y	421	IISASQEQIRLRTDDSSRRMWHIRRGESSRCGPYNLFNKRPILSNKYGOAYEVPEDYR	4800
D	b	481	QLDMKDVSFIANTITQGSMMGPFFINTNSTKVYVVASGEADVEMACPHLSGRHGRRGKR	5400
O	y	481	QLDMKDVSFIANTITQGSMMGPFFINTNSTKVYVVASGEADVEMACPHLSGRHGRRGKR	5400
D	b	541	HEEEDVHYEDEVKARKLSREAIVPVCPGVYFVSSGNENILLFAFGINAONNHENPLAGR	6000
O	y	541	HEEEDVHYEDEVKARKLSREAIVPVCPGVYFVSSGNENILLFAFGINAONNHENPLAGR	6000
D	b	601	ERNVLQOIEPQAMELAFAAPRKVEBELFNSDDESIFFPGRPQHQQSSRSSTKKQOQPLYSI	6600
O	y	601	ERNVLQOIEPQAMELAFAAPRKVEBELFNSDDESIFFPGRPQHQQSSRSSTKKQOQPLYSI	6600
D	b	661	LDFVGFF 666 	
O	y	661	LDFVGFF 666	
RESULT 2				
ID			W62828 standard; protein; 666 AA.	
AC			W62828;	
DT			27-OCT-1998 (first entry)	
DE			Macadamia integrifolia antimicrobial protein.	
KW			antimicrobial protein; infestation; control.	
OS			Macadamia integrifolia.	
FH			Key Location/Qualifiers	
FT			Peptide 1..28	
FT			/note= "signal peptide"	
FT			29..666	
FT			Protein /note= "mature protein"	
PN			MO9827805-A1.	
PD			02-JUL-1998.	
PF			22-DEC-1997; A00874.	
PR			20-DEC-1996; AD-004275.	
PA			(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.	
PI			Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;	
DR			WPI: 98-377279/32.	
DR			N-PSSD: V42310.	
PT			Novel anti-microbial protein from e.g. Macadamia integrifolia -	
PT			useful for controlling microbial infestations of plants or mammals	
PS			Claim 1; Page 34-36; 96pp; English.	
CC			The sequence is that of an antimicrobial protein which can	
CC			be used to control microbial infestations in plants and mammalian	
CC			animals.	
IQ			Sequence 666 AA;	

Query Match	96.6%	Score 4746:	DB 1,	Length 666;
Best Local Similarity	96.1%;	Pred. No.	0.00e+00;	
Matches	640; Conservative	Mismatches	10;	Indels 0; Gaps 0;
D b	1	MAINTSNLCSTLFLTSLTSTVSLASEEDRQREYECCKRCQMOLENSSGMRBCVSQCD	60	
O y	1	MAINTSNLCSTLFLTSLTSTVSLASEEDRQREYECCKRCQMOLETSGMARCVCSCD	60	
D b	61	KRFEEIDIMSYKDNDQEDPQTECCOCORCRQOESGPOROQCORRCKICEESEEYNOR	120	
O y	61	KRFEEIDIMSYKDNDQDPQTCOCOCORCRQOESGPOROQCORRCKICEESEEYNOR	120	

Dp	121	DPDOYECCKHCRREHREPHMHOTGQORERREKREKRRQKRYEEOQDEDEKTEER	180
Oy	121	DPQOQTECCOORCRHREPHMHOTCQORERREKREKRRQKRYEEOQDEDEKTEER	180
Dp	181	KEEDNKRDPQOREYEDCRRRCODEPFRQOQOLCRHEQORONGRGDMNPPRGSGRY	240
Oy	181	KEEDNKRDPQOREYEDCRRRCODEPFRQOQOLCRHEQORONGRGDMNPPRGSGRY	240
Dp	241	EEGEEOGSDNYRYDEDSLSRFRTEEGHLSYLENFGSKLLRALKNYLVLLEANPNA	300
Oy	241	EEGEEOGSDNYRYDEDSLSRFRTEEGHLSYLENFGSKLLRALKNYLVLLEANPNA	300
Dp	301	FVLPHLDADAILVIGRGALKIHHNDNESYNLECGDVIIRIPAGTFYLIRNDNNERL	360
Oy	301	FVLPHLDADAILVIGRGALKIHHNDNESYNLECGDVIIRIPAGTFYLIRNDNNERL	360
Dp	361	HAIAFLQITSTPGQYKEFFPAGGONPEPYLSTFSKELLEALNTQTEKLYGVEGOOREG	420
Oy	361	HAIAFLQITSTPGQYKEFFPAGGONPEPYLSTFSKELLEALNTQTEKLYGVEGOOREG	420
Dp	421	IIRASQEBIRBLTDDSDSRMHTRRGCESSRGPYLNFKNRPLYSNKYGAYEYKPEDR	480
Oy	421	IIRASQEBIRBLTDDSDSRMHTRRGCESSRGPYLNFKNRPLYSNKYGAYEYKPEDR	480
Dp	481	QLOQMDLSVFILANTQSGMMGPFFNTSTVYVVAASGEADVEAMACPHLSRHHGRGGGR	540
Oy	481	QLOQMDLSVFILANTQSGMMGPFFNTSTVYVVAASGEADVEAMACPHLSRHHGRGGGR	540
Dp	541	HEEEDVHYEEOVARLSKREALIVLAGHPVVFSSGNGENLLLEAFGINAQNHNEFLAGR	600
Oy	541	HEEEDVHYEEOVARLSKREALIVLAGHPVVFSSGNGENLLLEAFGINAQNHNEFLAGR	600
Dp	601	ERNVLOQIEPQAMLAAPAPKEYEESFNQODSIFPPGPRQHQOOSPRSTKOQPLVSI	660
Oy	601	ERNVLOQIEPQAMLAAPAPKEYEELFNSODESIFPPGPRQHQOOSPRSTKOQPLVSI	660
Dp	661	LDVYGF 666	
Oy	661	LDVYGF 666	

```

RESULT      3
ID          M62830 standard; Protein; 625 AA.
AC          M62830;
DE          27-OCT-1998 (first entry)
DE          Macadamia integrifolia antimicrobial protein.
DE          antimicrobial protein; infestation; control.
OS          Macadamia integrifolia.
FH          Key
FH          Location/Qualifiers
FT          Peptide
FT          1..28
FT          /note="signal peptide"
FT          Protein
FT          29..666
FT          /note="mature protein"
PN          WO9827805-A1.
PD          02-JUL-1998.
PF          22-DEC-1997: AU0874.
PR          20-DEC-1996: AU-004275.
PR          (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PA          Bower NI, Goulier KC, Green JL, Manners JM, Marcus JP;
PI          WPI: 98-377279/32.
DR          N-PSDB: V42316.
PT          Novel anti-microbial protein from e.g. Macadamia integrifolia -
PT          useful for controlling microbial infestations of plants or mammals
PS          Claim 1, Page 43-45, 36pp; English.
CC          The sequence is that of an antimicrobial protein which can
CC          be used to control microbial infestations in plants and mammalian
CC          animals.
SQ          Sequence      625 AA;
          4
Query Match      91.5%; Score 4496; DB 1; Length 625;
Best Local Similarity 96.6%; Pred. No. 0.00e+00;
Matches 604; Conservative 11; Mismatches 10; Indels 0; Gaps 0;

```


Db 36 DPKRYEDCRRRCWMDTRGQKEQOQSEESCKSQYGEKDQORHRRPDPQRCYEEQOECR 95
121 DPOQYEEQOECORRHETPRHMQTOCORCERY -EKEK -R-K-Q--QKRYE-QQ--R 170
Db 96 QOEERPOQOQACLRKFEDEQOQSOQROEQOQHOFQEPREKQOQYREKRYQEN 155
QY 171 EDEK-----YEER-MKEEDKRDPOQREYEDCRRCEQOEPQO--YOCORCREQOORH 223
Db 156 PMRGE-REEBAEET-EEGEQOSHNPFEHRSFOSRPREEHGNFVLRQARSHPTL 213
QY 224 GRGDDLINQRGSSGRYEEQEEQSDNPYTFDRSLSTRRTTEGHI SVLENTYGSKLL 283
Db 214 RGINEERLSTLEAMPNTFVLRPHCAEKIYLVNKGITLFTLHNEKESYNIYGVVYV 273
QY 284 RALKNYRLVLEANPAFVLPTHLDADAILLYTGGALMKMHRDRESYNLECGVIRI 343
Db 274 PGGSTYLLANODKKEKLIIVLRPNVNNQGFEEFPAGQSRQSLRAFSRILPAPN 333
QY 344 PGTITVLLINRDNRRRLHAKFLQITSTPGYKEFPAGQNEPYLSTFSKILPAAIN 403
Db 334 TRSEQDELFGGRSRRROQOGMFRKASQOETRALSOEAT-SPR-E--KSGE--RFAFN 387
QY 404 TQAEIRRGVLG-Q-QRE--GVIIASQEOIRRETRDSESRMHIRGESSRGPYN 456
Db 388 LLSQTPRYNQGREFEACRPFERQDINVTYSALQNGSIEFVHNSKATFVILTE 447
QY 457 LENKRPYLSNKYGAAYEVRPEYRQLOQMDVSVFANITOGSMGPFMTNSTKVVVVAS 516
Db 448 GNGYAMVSPHLPROSSYEEDEEEOEEOEERSQYKIRSRRLRGDIFYVPANF 507
QY 517 GENDVEMACPRL--SG-RHGRGGRK-HEEED--V-HYEOVKARLSKREIIVPVGH 568
Db 508 PVTYVASQONQNTMRTGFGLYNINPDNQIRFVAGKINH-V-ROMSQAKELAFGVSSRL 566
QY 569 PVTYVSSGNENLLFAFGINAQ-N-H-EN-FLAGRENVLIQIERPQAMELAFAPRKE 623
Db 567 VDEIENSNOESTEVVS-RQRORAS 589
QY 624 VELEFNSODESIFFPGRHOQOOS 647

RESULT 6
ID W62831 standard; Protein: 525 AA.
AC W62831;
DT 27-OCT-1998 (first entry)
DE Theobroma cacao antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Theobroma cacao.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 47-49; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
be used to control microbial infestations in plants and mammalian
CC animals.
CC Sequence 525 AA;

Query Match 24.3%; Score 1193; DB 1; Length 525;
Best Local Similarity 40.6%; Pred. No. 9, 54e-93;
Matches 195; Conservative 122; Mismatches 140; Indels 23; Gaps 17;

Db 33 QYERDPROQYEQCORCESE-ATEREDE-QCE--QRCERE-YKQOQROEE--E-LOQO 84
QY 133 RHETPRHMH-QYCOQRCERYEKEKRRKQKRYEQOQREDEKYEERMKEDKRDPOQRE 193
Db 85 YQOCCGROEQOQOQOQOQOQCKCKEYQYKEBERG-EHENTYNNHKKNSSEEGQORNNP 143

QY 194 YEDCRRRC-EOQE-PROQYOCORCREQORHGRGDDLINPORGSGRYEEQOQSDNP 251
Db 144 YMFPRRSQOTFRDEEGNFKLQRFAPENSPLKGINYLRLAMEFANPTEFLPHCAE 203
QY 252 YTFDER-SLSTFRTEEGHISYLENFYGRSKLLRALKNRLVLEANPAFVLPTHLDAD 310
Db 204 AIFYVTNGKGTTFTEHNEKESYNQRTGVVSVPPAGSTVYVVSODNEKLTIAVLTALPVN 263
QY 311 ALLVTGSGALMKMHRDRESYNLECGVIRIPAGTIFYLLINRDNRRRLHAKFLQIT 370
Db 264 SFGKTELEFPAGNNKPESTYGAFTVLETVFNTOREKLEETLEBQROKROQOQGMFR 323
QY 371 TPGQYKEFPAGQNEPYLSTFSKILEALMTOAEIRGLVGOQREGVIIISASQEOIR 430
Db 324 KAKPRQIRASQOASPRRGERELAINLSQSPYVSNOQNEFEACPEDSDPONMVA 383
QY 431 ELTRDD--SESRKMHIRKGESSRGPYNLFNKRPYLSNKYGAAYEVRPEYRQLOQMDVS 488
Db 384 VSAFLNQAIFVPHYNSKATPEVFTDGYGAQMACPHLSROSQOSGRODRREQEE 443
QY 489 VLIANTOGSMGPFMTNSTKVVVVASGADVEMACPRLSGR-HGGRG--GKR-HEE 544
Db 444 SEETFGGEQYKAPLSDGVFVAPAGNAVTFPASKDOPLANAVAGLANQNNQIFLAGR 503
QY 545 --EDV--HYEOVKARLSKREIIVPVGHVYVSSGNENLLFAFGINAQNNHENTFLAGR 600

RESULT 7
ID W62835 standard; Protein: 593 AA.
AC W62835;
DT 27-OCT-1998 (first entry)
DE Zea mays antimicrobial protein.
KW antimicrobial protein; infestation; control.
OS Zea mays.
PN WO9827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-004275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;
DR WPI: 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 58-60; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
be used to control microbial infestations in plants and mammalian
CC animals.
CC Sequence 593 AA;

Query Match 22.5%; Score 1106; DB 1; Length 593;
Best Local Similarity 40.8%; Pred. No. 6, 29e-85;
Matches 204; Conservative 112; Mismatches 155; Indels 29; Gaps 22;

Db 25 EDDNNHHHGHKSGQVCARCEDR-PWHORPCLQECREERKEKQERSRHADRSGECS 83
QY 182 EEDNKRDPQOQREYEDCRRCEQOEPQOY-QCORCREQOQO--HGRGDDLINPORG-GS 237
Db 84 SEDERQEKEROKDRPVDFPDRSRVVRVSEQSLVLRPFDEVSRLRGIRYRAVL 143
QY 238 GRYE-EGE-EKQSDN-PYFDRSLSTFRTEEGHISYLENFYGRSKLLRALKNRLVLL 294
Db 144 EANPRSFVPSHTAHACIYVAEGVYTTIENGERRSYTIKQGHVFAPAVATYTLANT 203
QY 295 EANPAFVLPVLRHLDADAILLYTGGALMKMHRDRESYNLECGVIRIPAGTIFYLLNR 354
Db 204 DGRKLVYTKLHITSVGEFQFFPGCGNPFSLFSFSISQAPAYKSSDRLERLFG 263
QY 335 DNNEERLHAKFLQITSTPGYKEFPAGQNEPYLSTFSKILEALMTOAEIRGLVIG 414
Db 264 RHGQDKGIVATQTEQTELRHNSGEGHGHVPLPFGES-RGPYSLDORPSIANQHG 322
QY 415 QO-RE-GVIIASQEOIRRETRDSES-R--RWHIRKGESSRGPYNLFNKRPYLSNKYG 469

QY 367 Q-TISTPGQYKEFFPAGGONPEYLSTFSKEITLALNTQAEIRLGVLGQOREGVIISAS 425
DB 287 REQVRLAAPTAKSSM-WF--GGES-KPOFNIFSKRPTIISNGYGLTLEVPPDDEKSMQ 342
QY 426 QEOIRELTRDSESRMRHIRRGESSRGPYNLFKRRLYKNKQGALEVAPEDY-RQ-LQ 463
DB 343 RLNLMLTFNITORSMTIYHNSHATIALYIDGRHLQISCPHMSR-SSH--SK-HDK 398
QY 484 DMQVYFIANTIGOSMGPFENSTKVVVAVASGEADVENACPHLSRGHGRGGRKHEE 543
DB 399 SSPS-YHRISSDLKPGVAVFVPPCHPPTVIAKNENLMICFVNADNK-KTTFAGDK- 455
QY 544 EEDVHTEQVAFARLSKRAPIVAVGHPVAVSSGNNLLLEFAGINAGNHENF-LAGRER 602
DB 456 NIVSLLDNVAKELAFNYPSEMVNCF 481
QY 603 NVLQIIEPQAMELAFAPRKEVELEF 628
RESULT 10
ID W62838 standard; protein; 605 AA.
AC W62838;
DT 27-OCT-1998 (first entry)
DE Glycine max antimicrobial protein.
KW antimicrobial protein; Infestation; control.
OS Glycine max.
PN W09827805-A1.
PD 02-JUL-1998.
PF 22-DEC-1997; AU0874.
PR 20-DEC-1996; AU-0042275.
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
PI Bremer NJ, Goulter KC, Green JL, Manners JM, Marcus JP;
WPI; 98-377279/32.
PT Novel anti-microbial protein from e.g. Macadamia integrifolia -
useful for controlling microbial infestations of plants or mammals
PS Claim 1; Page 63-65; 96pp; English.
CC The sequence is that of an antimicrobial protein which can
be used to control microbial infestations in plants and mammalian
CC animals.
SQ Sequence 605 AA;
Query Match 20.0%; Score 981; DB 1; Length 605;
Best Local Similarity 32.8%; Pred. No. 9,98e-74;
Matches 194; Conservative 160; Mismatches 195; Indels 43; Gaps 32.
DB 38 CLOSENSESDSYRNOA-CHARCNILKVECEGEIIPRPPPOHPR-EPQOPEKEE 95
QY 85 CQRRCQOEGSPHQOQCQRCKEICEEBEYRQNDPQO-YEQOERQRIETPRHM 143
DB 96 DEDEQPRPIPPPPROBE--EHOOREOE-WP-R-KEE--KREKSEED-EDDEE 147
QY 144 QYQOQCEKERYEKKRQOYKREEOQREDEKEERKEEDNRDPQOQREYEDCRRCEQ 203
DB 148 QDERO-FPPFRPHOKEREDEDEEORESEES-EDSELRRHKNKNFELGSRFEI 205
QY 204 QERPOQOCORCEREOQRGGLINPORGSGRYESEE-EKQSD-NPYTDESLST 261
DB 206 LFNQYGRITVLOFNOFQSPOLNDYRILEFNSKPNITLLPNHADADVLITLNGTAT 265
QY 262 RFTLEGHSIVLENFGRSKLLALKNRYVLEANPNVFLPTHADADILLVGTGRGA 321
DB 266 LSLVNNDDRDSTYLGSDALRPVSGTYYVVDNNENLRLTLAIPVKNKGRFESFIS 325
QY 322 LKMIHRDNRESYMLEGCDVIRIPAGTTFYLINDNNEERLHIAFLQTIISPGQKEFFPA 381
DB 326 STAAQSYLOGEFRNLLEASDYTKFEEINKVLFSSREGOQOQSGORLOESYVIAISQQR 385
QY 382 GCGNPEPILSTFSKEITLALNTQ-AE--R-L--R--GV-L-GQOR--EGVTIASQEQIR 430
DB 386 ALSKRAKSSSRK-TI--SSSD-K-PNLRSRDPYISNKLKFEETPEKPNQDLIDIL 440
QY 431 ELT-RDSESRMRHIRRGESSRGPYNLFKRRLYKNKQGALEVAPEDYRQLQDMQVAV 489

DB 441 SYVDNEGALLPHFNSKAIVILVINEGDANIELV-G-LKEQOEOQ--QEOPLEVR- 494
QY 490 FIANTIGOSMGGEFFNTRSTKYVVAVASGEADVENACPHLSRGHGRGGRKHEEEDVAV 549
DB 495 -KTRALSODIDIVITAGIPVY-VNAT-S-NLNFPAIGIAENNRFLAGSDONVYSQIP 551
QY 550 EOYKATLSKREALIVVGVHPVAVSSGNNLLLEFAGINAGNHENFLAGRENNVLOQIE 609
DB 552 SOVQELAFPGSAQAVELKLNQRESYFYDAOPKKKEGKGRG--PLSSIL 601
QY 610 PQAMELAFAPRKEVELELNSODESIFFPGRHOQOQSSRSTKQOOPVYSIL 661
RESULT 11
ID W90340 standard; protein; 444 AA.
AC W90340;
DT 24-MAY-1999 (first entry)
DE G. max truncated SBP1 protein.
KW SBP1; sucrose binding protein; SBP2; sucrose uptake; transgenic plant;
seed; carbohydrate content; soybean.
OS Glycine max.
PN W09853086-A1.
PD 26-NOV-1998.
PF 21-MAY-1998; U10465.
PR 22-MAY-1997; US-047568.
PA (UNIT) UNIV WASHINGTON STATE RES FOUND.
PI Chao WS, Grimes HD;
WPI: 99-070155/06.
PT New modified plant sucrose binding proteins - used to develop
transgenic plants which can have enhanced or decreased sucrose
uptake activity in developing seeds
PS Claim 7; Page 36-37; 58pp; English.
CC This sequence represents a novel sucrose binding protein, SBP1 isolated
from Glycine max. This protein is used in a method resulting in the
production of a modified plant sucrose binding protein (SBP) which has a
modified amino acid sequence compared to a corresponding wild-type SBP,
and where expression of the modified SBP in a yeast assay system confers
enhanced sucrose uptake compared to the corresponding wild-type SBP.
CC The products of the invention can be used for producing transgenic plants
CC which have modified sucrose uptake activity, particularly in developing
CC seeds. Enhanced sucrose uptake activity in developing seeds may be
desirable where it is an advantage to increase the carbohydrate content
of the seed (e.g. where the seed is the primary plant material harvested,
such as soybean). In contrast, decreased sucrose uptake activity in
CC seeds might be desirable where the vegetative material of the plant is
harvested. The SBP regulatory regions confer specific or enhanced
CC expression in developing seeds and so may be used to express any
transgene in developing seeds.
SQ Sequence 444 AA;
Query Match 19.0%; Score 933; DB 1; Length 444;
Best Local Similarity 37.8%; Pred. No. 1,94e-69;
Matches 154; Conservative 110; Mismatches 126; Indels 17; Gaps 15;
DB 49 QOQQTGEGDRVC-LQ-SCDRIYRNMKQERKQIOETREKKEESREREDEQOHEOD 106
QY 190 QOREY-EDCRRCEQOERPOQOCORCEREOQRGGLINPORGSGRYESEEKRS 248
DB 107 EMPYIEEKDEETVEDEGRIYVKKFTESKLLQSTENFRLAILARATYSPNHF 166
QY 249 DNPYTFDE-RSLSTRFEEGHSIVLENFYGRSKLLRLKKNRYLLEANPNVFLPTHL 307
DB 167 DEVEVFENKGRVAVGLVSESETEKITLEPGMIHIPGTPIYVNRBENKLFAMLI 226
QY 308 DADAILIYVGKALKMIRHNRRESYNLECGDIVIRIPAGTTFYLINDNNEERLHIAK-L 366
DB 227 PVSVSTPGKFEFFAPGGRDPESVLSAFSWNLQAALQTPKGLTENVEDQONEGSIPIRIS 286
QY 367 Q-TISTPGQYKEFFPAGGONPEYLSTFSKEITLALNTQAEIRLGVLGQOREGVIISAS 425
DB 287 REQVRLAAPTAKSSM-WF--GGES-KPOFNIFSKRPTIISNGYGLTLEVPPDDEKSMQ 342
QY 426 QEOIRELTRDSESRMRHIRRGESSRGPYNLFKRRLYKNKQGALEVAPEDY-RQ-LQ 483

Dd	363	RLLMLFETNTIQSGMSPFHNSATKTKALYIDGCHQICPMHSSP--SK-HDK	398
Oy	484	DDMDVSFLANTIIQGSMGPFFINTSTKYVVAASGEADVEAMCPHLGHGRGGRRKKHE	543
Dd	399	SSPS-YHRISSDLKGCWAVEVPYPHPFTVALSNKENLIMICEVNAR	444
Oy	544	EEDVAHQVKRLRSKEAIVPVGHPVFVSSGNENLLFAFGINAQ	590
<hr/>			
RESULT	12		
ID	W22150	standard; Protein; 626 AA.	
AC	M22150:		
DT	29-DEC-1997	(first entry)	
DE	Peanut allergen Ara hi.		
KW	Peanut seed storage protein; allergen; allergy; hypersensitivity;		
KW	vaccine; anaphylactic shock; immunotherapy; therapy;		
KW	monoclonal antibody; ELISA; analysis; Ara hi.		
OS	Arahis hypogaea strain Florunner.		
FH	Key	Location/Qualifiers	
FT	Peptide	1..22	
FT	/label= Sig-peptide		
FT	Protein	23..626	
FT	/label= Mat-protein		
FT	Modified_site	521..523	
FT	/note= "N-glycosylation site"		
PN	WO9724139-A1.		
PD	10-JUL-1997.		
PF	23-SEP-1996; U15222.		
PR	04-MAR-1996; US-610424.		
PR	29-DEC-1995; US-009455.		
PA	(UOAR-) UNIV ARKANSAAS.		
PI	Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;		
DR	WPJ: 97-363453/33.		
DR	N-PDB: T76613.		
PT	Peanut allergens Ara hi and Ara hII - used for vaccination and in		
PT	two-site monoclonal antibody based ELISA		
PS	Claim 31; Page 172; 33pp; English.		
CC	This polypeptide comprises major peanut allergen Ara hi (W22149).		
CC	Its sequence was deduced from cDNA clone P41b (T76613), isolated		
CC	from peanut seed cDNA using a primer (see T76616) based on an		
CC	isolated Ara hi peptide (see W24206). The sequence shows		
CC	significant homology with the vicilin family of seed storage		
CC	proteins of other legumes. The allergen is recognised by serum		
CC	IgE from a large proportion of individuals with peanut		
CC	hypersensitivity. Ara hi and Ara hII (see W24164) can be used to		
CC	raise monoclonal antibodies which are used in a specific two-site		
CC	MAb ELISA for the detection of Ara hi or Ara hII (claimed). IgE-		
CC	binding Ara hi antigen epitopes (see W24165-87) may be used in		
CC	cc vaccines to protect against allergic reactions to peanut allergens,		
CC	e.g. anaphylactic shock.		
CC	Sequence 626 AA;		
<hr/>			
Query Match	19.0%; Score 933; DB 1; Length 626;		
Best Local Similarity	33.6%; Pred. No. 1.94e-69;		
Matches 195; Conservative 156; Mismatches 188; Indels 42; Gaps 30;			
Dd	38	CAQRCLSCGOEPDDLKKAACESCKT-LFTDPRCVTYDPRGHTTTNQSRPG-EFRRCR	95
Oy	102	CQRCEKEECSEEEENRRDQQOYEQCQECQHETEPR-HMQTCOORERREYKEKRK	160
Dd	96	OPGVDDDDRPPRPREGGR-WGAPGRE-FERE-ED-WRO-PREDMWRRPSH-OOP-R-KI	147
Oy	161	QOKRREQRDDEKEYERMEKNRKDPQOREFEDCRRREQOEPRQOYTQCARCEEQ	220
Dd	148	RPEGEGE---QEWGTPEGSVHR-ETSNNDFYFPSSRFSTRYGNGRIYLORFORS	203
Oy	221	ROHGSGCDLIPNOGSGSRYEEGEEKOSDNMYFEDESLSTRFTEGHSTYLENFYGRS	280
Dd	204	ROFOUNLRHRTVOLEAKNTVLTLKKHADADVIITVOOGAQAVTYVANNNKRSFYLDGHA	263
Oy	281	KLRRLKRVLTVLLEANNAFVLPTHIDADDILTVTGGRGLKAIHHNDNRSYILECGDV	340

```

Db      264  LTRSGGLSYLNRHNDONLRKATSMNVTPGCGEDFEPASSDQSYLQGSNLTLEA 323
Qy      341  IRIPIGTFYLLNRRNNERLIRAKFLQTIISPGQKFEFFPAGGQNPPEYISTFSEILEA 400
Db      324  AFNAFENFIRPRLLEENNGGEGOEERQGRMSTRSSENNEGVIVKSKHEVELTKHAKSV 383
Qy      401  ALNTQAEERKGL-----GQORE-GVLI-SA--SQE-Q--IRELTRPD-SE-SRR-WHI 444
Db      384  SKKSGSEGDITNPILNRGEFDDLSNNEGKLFYVKKPDKKNPOLDDMLMLTCVEIKEGAL 443
Qy      445  -RRGGE--SS-RGPYNLFRKRLPYNSKNGQAYEVKPED-YHQLQMDVSVFIANTQSGM 499
Db      444  MLEPHNSKAMVIVVYVVKGTGNLELVAVAKEDQQRGRREEDEDEDEEGSSREYRRTIAR 503
Qy      500  MGPPEFTSTKVVVAVASGEADVEMACPHLSGRHGRGRGRKHE--EEDDVHYEYVK--AR 555
Db      504  LKEGVFTMPAAHPVA-INASSE-LHLGFGIMNENNRIFLAGDKVDVQIDQIKOAKDL 561
Qy      556  LSKRAIIVPVGHVPEVYVSGENELLEPAFGIMQNNHENFLAGKERRNVLDQIEPQAMEL 615
Db      562  AEPGSGQEVEKLIRKQKSEHFVSARPOSOSPPSSPEKESP 602
Qy      616  AFAAPRKEVEELFNQDSEITFPFGPRHQDQSSNSTKQGP 656

RESULT 13
ID      W22149 standard; Protein:4614 AA.
AC      W22149;
DT      29-DEC-1997 (first entry)
DE      Peanut allergen Ara hi.
KW      Peanut; seed storage protein; allergen; allergy; hypersensitivity;
KW      vaccine; anaphylactic shock; immunotherapy; therapy;
KW      monoclonal antibody; ELISA; analysis; Ara hi.
OS      Arachis hypogaea strain Florunner.
FH      Key
FT      Peptide      location/Qualifiers
FT      1..22
FT      /label= Sig-peptide
FT      Protein      23..614
FT      /label= Mat-protein
FT      Modified_site 521..523
FT      /note= "N-glycosylation site"
FN      W09724139-A1.
PD      10-JUL-1997.
PF      23-SEP-1996; U15222.
PR      04-MAR-1996; US-610424.
PR      29-DEC-1995; US-009455.
PA      (UYAR-) UNITV ARKANASAS.
PI      Bannan GA, Burks AW, Cockrell G, Helm RM, Stanley JS;
DR      WPI: 97-363453/33.
DR      N-PSDB: T76612.
PT      Peanut allergens Ara hi and Ara hII - used for vaccination and in
PT      two-site monoclonal antibody based ELISA
PS      Claim 31, Page 169, 354pp; English.
CC      This polypeptide comprises major peanut allergen Ara hi (W22149).
CC      Its sequence was deduced from cDNA clone P17 (T76612), isolated
CC      from peanut seed cDNA using a primer (see T76616) based on an
CC      isolated Ara hi peptide (see W24206). The sequence shows
CC      significant homology with the vicilin family of seed storage
CC      proteins of other legumes. The allergen is recognised by serum
CC      IgE from a large proportion of individuals with peanut
CC      hypersensitivity. Ara hi and Ara hII (see W24164) can be used to
CC      raise monoclonal antibodies which are used in a specific two-site
CC      Mab ELISA for the detection of Ara hi or Ara hII (claimed). IgE-
CC      binding Ara hi antigen epitopes (see W24165-87) may be used in
CC      vaccines to protect against allergic reactions to peanut allergens,
CC      e.g. anaphylactic shock.
SQ      Sequence 614 AA;

Query Match      18.8%; Score 922; DB 1; Length 614;
Match Local Similarity 34.0%; Pred. No. 1.86e-68;
Matches 192; Conservative 158; Mismatches 173; Indels 42; Gaps 30.

45 QOEPPDLKQKCESRCTLEYDPCVYDTGAT-NORHPGGERRG-R-QEDYDDBDR-Q 100

```


Matches 146; Conservative 105; Mismatches 117; Indels 18; Gaps 14;

```

Db      35 EDPETVTCCKHOCQOO--ROYTESDKRTCLQCCDSMKOEKREKOVEEETREKEEHOEHHE 92
      190 QCRREYEDCRRCEQOEPRQOYCQRR-CREQ-QR-QHGRGDLINPQRGSGRYEEG--EE 245
Db      93 EEDENPYVEEDKDFSTRVETEGGSIIVLKKFTEKSKLLQGIENFRLAILEARAHFVSP 152
      246 KQSDNPYYFDE-RSLSTRFRTTEGHISVLENFYGRSKILRALKNYRLVYLEANPNAFVLP 304
Db      153 RHFDSEVYFNINIKRAVLGLVRESETEKITLEPQDMIHIPACTPLYIVRDENEKILLAM 212
      305 THLDADAILVTGGKALKMHIRDNRESYNLECGDVIRIPAGTTFYLINRDNNERLHIAK 364
Db      213 LHIPTVTPGKEFEFFPGGROPESVLSAFSMNVLOALQTPKGLERLFPNOQNEGSIFKI 272
      365 FLQITISTPGQYKEFFPAGQNPPEYLISTFSKILEALNTQAERLRGVILGQOREGYIISA 424
Db      273 SREVRALAPTKKSSW-WPF--GSES-KAQENIFSKRPTFSNGYGRLTVEGPDDEKSWLQ 328
      425 SQEQIRELTRDDESRRWHIRRGESSRGPYNLFNKRPLYSNKYGQAYEVKPEDEYRQ-LQ 483
Db      329 RUNLMLTFNITQRMSTIHYNSHATKIALVMDGRGHQISCPHMSR--S--DSK-HDK 383
      484 DMDVSVFIANITQSGSMGPFENFTRSTKVVVVASGEADVEMACPHLSGRHGRGGRKHEE 543
Db      384 SSPPS-YHRISADLKPGWVFPVPPGHP 408
      544 EEDVHYEQVKARLSKREAIYVPVGHHP 569
  
```

Search completed: Sat May 13 08:09:07 2000
 Job time : 29 secs.

